Robust OED of Infinite-Dimensional Nonlinear Bayesian Inversion

Abhijit Chowdhary [1] Ahmed Attia [2] Alen Alexanderian [1]

[1]North Carolina State University [2]Argonne National Laboratory

Main Ideas

Traditional optimal experimental design (OED) methods are blind to misspecification in the hyperparameters of the inverse problem. Robust OED methods aim to address this, but have been

- limited to linear inverse problems, or
- limited to low-dimensional problems.

In order to solve robust OED problems for **infinite-dimensional nonlinear** Bayesian inverse problems governed by PDEs, we propose a formulation with

- a **budget-constrained probabilistic encoding** of the sensor locations, and
- adjoint-based **eigenvalue sensitivity** techniques for differentiation.

Worst-Case Robust OED Formulation

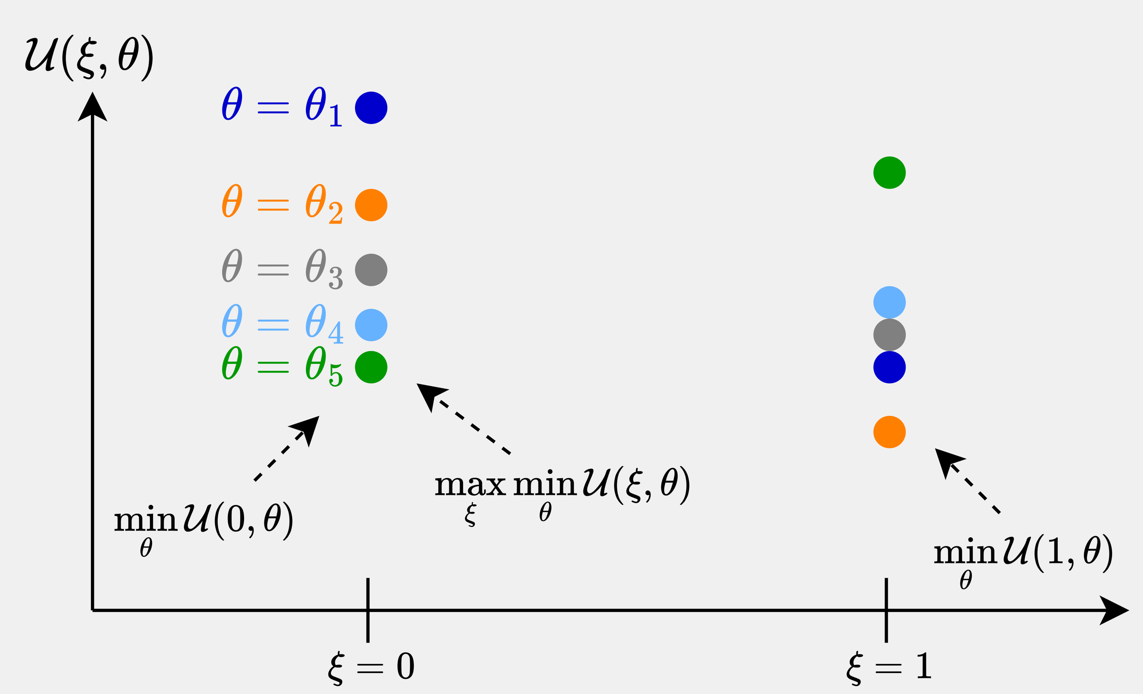
Consider the set of candidate sensor locations $\mathcal{S} = \{s_1, s_2, \dots, s_{N_d}\}$, and let $N_b \ll N_d$ be the budget constraint on the number of sensors. Let $\boldsymbol{\xi} \in \{0,1\}^{N_d}$ be a binary encoding of the observational configuration such that ξ_i determines whether s_i is active, and let $\boldsymbol{\theta} \in \Theta$ be the uncertain parameter. The ROED problem is defined as the optimization problem

$$\max_{\boldsymbol{\xi} \in \mathcal{S}(N_b)} \min_{\boldsymbol{\theta} \in \Theta} \mathcal{U}(\boldsymbol{\xi}, \boldsymbol{\theta}),$$

where

$$\mathcal{S}(N_b) = \left\{ \boldsymbol{\xi} \in \{0,1\}^{N_d} : |\boldsymbol{\xi}| = N_b \right\},$$

and the utility (objective) $\mathcal{U}$ is chosen to quantify the quality of the design.



Budget-Constrained Probabilistic Robust OED

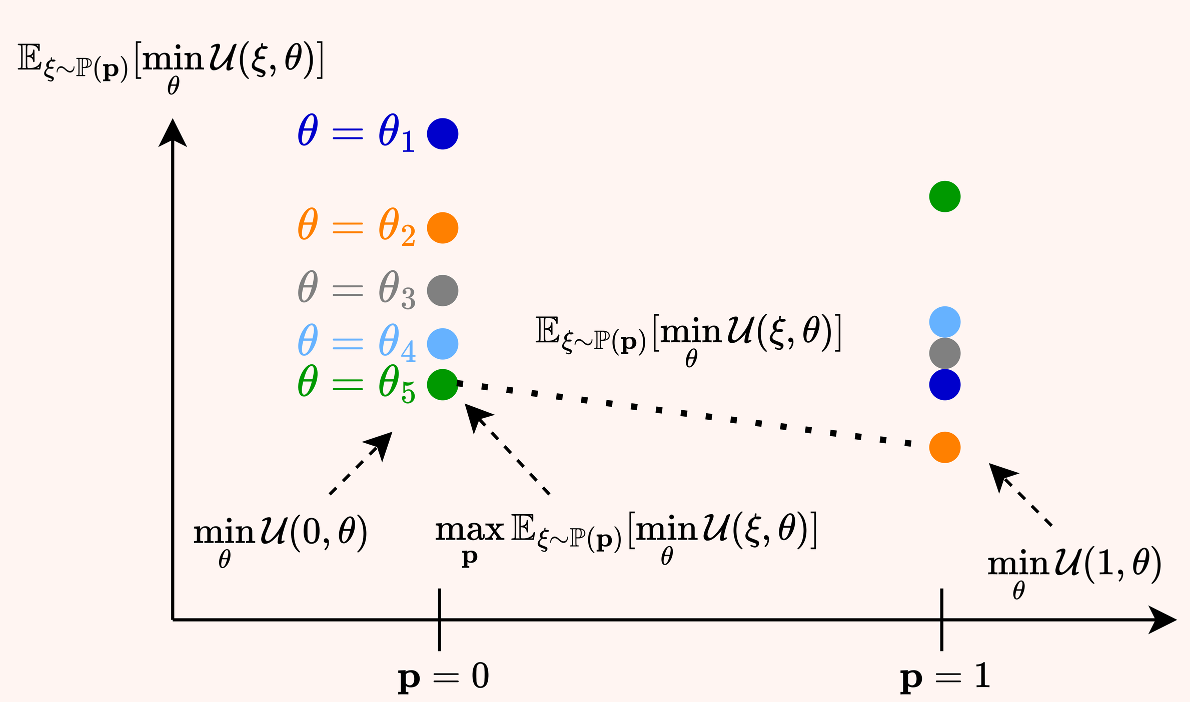
Assume that $\boldsymbol{\xi}$ is a random variable endowed with the **conditional Bernoulli distribution** $\mathbb{P}(\boldsymbol{\xi}|\mathbf{p}, |\boldsymbol{\xi}| = N_b)$. Then, the **budget-constrained** probabilistic robust OED problem replaces the classical robust OED formulation with the following policy optimization problem:

$$\max_{\mathbf{p} \in [0,1]^{N_d}} \mathfrak{U}(\mathbf{p}) := \mathbb{E}_{\boldsymbol{\xi} \sim \mathbb{P}(\boldsymbol{\xi}|\mathbf{p}, |\boldsymbol{\xi}| = N_b)} \left[\min_{\boldsymbol{\theta} \in \Theta} \mathcal{U}(\boldsymbol{\xi}, \boldsymbol{\theta}) \right].$$

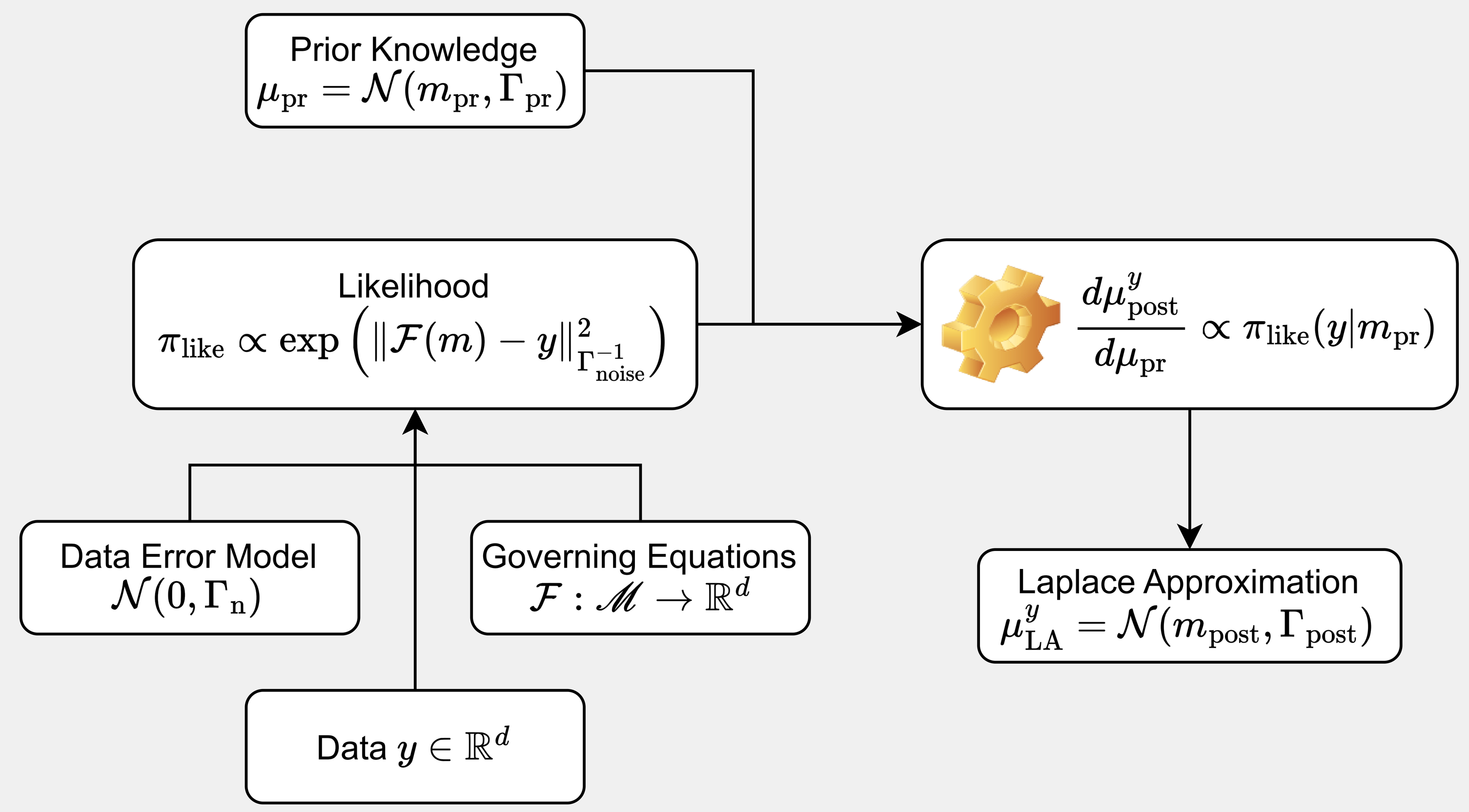
We denote $\mathfrak{U}$ as the stochastic objective. Furthermore,

$$\boldsymbol{\nabla}_{\mathbf{p}} \mathfrak{U}(\mathbf{p}) \approx \frac{1}{N_{ens}} \sum_{k=1}^{N_{ens}} \left[\min_{\boldsymbol{\theta} \in \Theta} \mathcal{U}(\boldsymbol{\xi}[k], \boldsymbol{\theta}) \ \boldsymbol{\nabla}_{\mathbf{p}} \log \mathbb{P}(\boldsymbol{\xi}[k] | \mathbf{p}, |\boldsymbol{\xi}| = N_b) \right].$$

Critically, this does not require design gradients of $\mathcal{U}$! Nevertheless, does require $\mathcal{U}_{\boldsymbol{\theta}}$.



PDE-Constrained Nonlinear Bayesian Inverse Problems



The forward model is governed by the PDE: Given $m \in \mathscr{M}$, find $u \in \mathscr{U}$ such that $a(u, m, p) = 0$ for all $p \in \mathscr{V}$. Furthermore, we employ the **Laplace approximation** to the posterior $\mu^{\mathbf{y}}_{LA}$.

Sample Averaged Expected Information Gain (EIG)

We employ a sample-averaged approximation of the EIG as our utility:

$$\begin{aligned} \overline{D_{KL}} &:= \mathbb{E}_{\mathbf{y}} \left[D_{KL}\left(\mu^{\mathbf{y}}_{post} \,\|\, \mu_{pr} \right) \right] \\ &\approx \frac{1}{N_{SAA}} \sum_{i=1}^{N_{SAA}} D_{KL}(\mu^{\mathbf{y}_i}_{post} \,\|\, \mu_{pr}) \\ &\approx \frac{1}{N_{SAA}} \sum_{i=1}^{N_{SAA}} \frac{1}{2} \left[\log\det\left(\mathcal{I} + \widetilde{\mathcal{H}}^i_m \right) - \operatorname{tr}\left(\widetilde{\mathcal{H}}^i_m [\mathcal{I} + \widetilde{\mathcal{H}}^i_m]^{-1} \right) + \|m^i_{post} - m_{pr}\|^2_{\mathcal{C}^{-1}_{pr}} \right] \end{aligned}$$

where we have assumed a **Laplace approximation** to the posterior measure and $\widetilde{\mathcal{H}}_m = \mathcal{C}^{1/2}_{pr} \mathcal{H}_m \mathcal{C}^{1/2}_{pr}$. Since $\widetilde{\mathcal{H}}_m$ is typically low-rank, we furthermore write:

$$\overline{D^{(r)}_{KL}}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \frac{1}{2N_{SAA}} \sum_{i=1}^{N_{SAA}} \left[\sum_{n=1}^{r} \left[\log\left(1 + \lambda^i_n(\boldsymbol{\xi}, \boldsymbol{\theta})\right) - \frac{\lambda^i_n(\boldsymbol{\xi}, \boldsymbol{\theta})}{1 + \lambda^i_n(\boldsymbol{\xi}, \boldsymbol{\theta})} \right] + \|m^i_{post}(\boldsymbol{\xi}, \boldsymbol{\theta}) - m_{pr}\|^2_{\mathcal{C}^{-1}_{pr}} \right].$$

Adjoint-Based Eigenvalue Sensitivity Techniques

The forward model is governed by the PDE: Given $m \in \mathscr{M}$, find $u \in \mathscr{U}$ such that $a(u, m, p) = 0$ for all $p \in \mathscr{V}$.

It can be demonstrated that $(\mathcal{H}_m, \{\lambda_n, \psi_n\}^r_{n=1})$ obey the eigenproblem constraints

$$\begin{aligned} \langle \phi, \mathcal{H}_m \psi_n \rangle &= \lambda_n \langle \phi, \psi_n \rangle_{\mathcal{C}^{-1}_{pr}}, & \forall \phi \in \mathscr{V}, \forall n = 1, \dots, r, \\ \langle \psi_n, \psi_n \rangle_{\mathcal{C}^{-1}_{pr}} &= 1, & \forall n = 1, \dots, r, \end{aligned}$$

and $\mathcal{H}_m$ has the following system dictating its action

$$\mathcal{H}_m(m)(\psi_n, \phi) = \langle \phi, a_{mp}(u, m, p)\hat{p} \rangle,$$

with state and adjoint constraints

$$\begin{aligned} \langle \tilde{p}, a_p(u, m, p) \rangle &= 0, & \forall \tilde{p} \in \mathscr{V}, \\ \langle \tilde{u}, a_u(u, m, p) \rangle + \left\langle \tilde{u}, \boldsymbol{\mathcal{Q}}^* \widehat{\boldsymbol{\Gamma}}^\dagger_n(\boldsymbol{\xi}, \boldsymbol{\theta})(\mathbf{y} - \boldsymbol{\mathcal{Q}} u) \right\rangle &= 0, & \forall \tilde{u} \in \mathscr{U}, \end{aligned}$$

and incremental state and adjoint constraints for $n = 1, \dots, r$:

$$\begin{aligned} \langle \tilde{p}, a_{pu}(u, m, p)\hat{u}_n \rangle + \langle \tilde{p}, a_{pm}(u, m, p)\psi_n \rangle &= 0, & \forall \tilde{p} \in \mathscr{V}, \\ \langle \tilde{u}, a_{up}(u, m, p)\hat{p}_n \rangle + \left\langle \tilde{u}, \boldsymbol{\mathcal{Q}}^* \widehat{\boldsymbol{\Gamma}}^\dagger_n(\boldsymbol{\xi}, \boldsymbol{\theta}) \boldsymbol{\mathcal{Q}} \hat{u}_n \right\rangle &= 0, & \forall \tilde{u} \in \mathscr{U}. \end{aligned}$$

Hence, after fixing the MAP estimate in $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$, we define the utility

$$\mathcal{U} := \frac{1}{2N_{SAA}} \sum_{i=1}^{N_{SAA}} \left[\sum_{n=1}^{r} \left[\log\left(1 + \lambda^i_n(\boldsymbol{\xi}, \boldsymbol{\theta})\right) - \frac{\lambda^i_n(\boldsymbol{\xi}, \boldsymbol{\theta})}{1 + \lambda^i_n(\boldsymbol{\xi}, \boldsymbol{\theta})} \right] + \|m^i_{post}(\boldsymbol{\xi}^{all}, \overline{\boldsymbol{\theta}}) - m_{pr}\|^2_{\mathcal{C}^{-1}_{pr}} \right].$$

We can differentiate through this using a formal Lagrangian approach.

Numerical Experiments

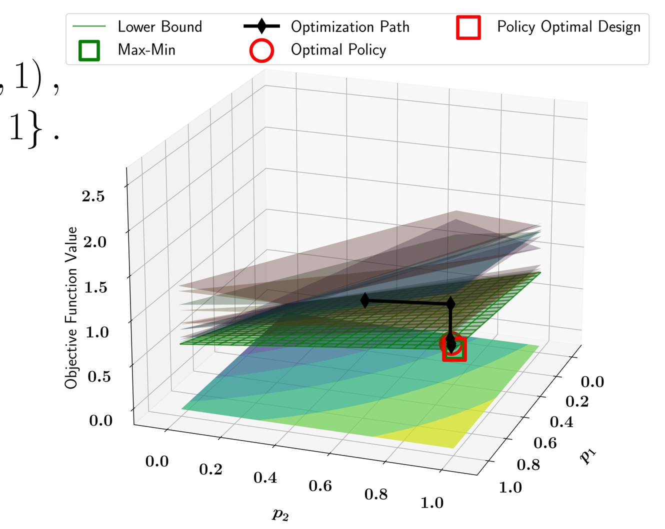
$$\begin{aligned} -\nabla \cdot (\exp(m)\nabla u) &= 0 & &\text{in } \Omega := (0,1)^2, \\ \exp(m)\nabla u \cdot \mathbf{n} &= 0 & &\text{on } \Gamma_N := \{0,1\} \times (0,1), \\ u &= g & &\text{on } \Gamma_D := (0,1) \times \{0,1\}. \end{aligned}$$

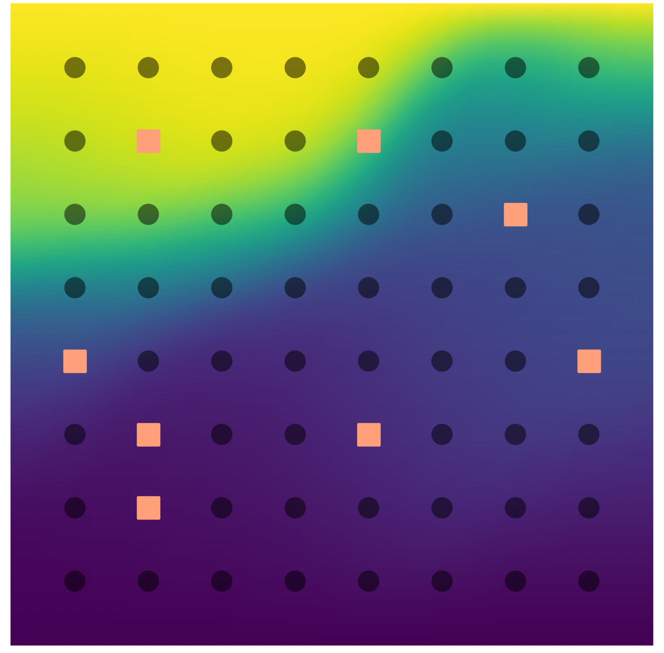
and

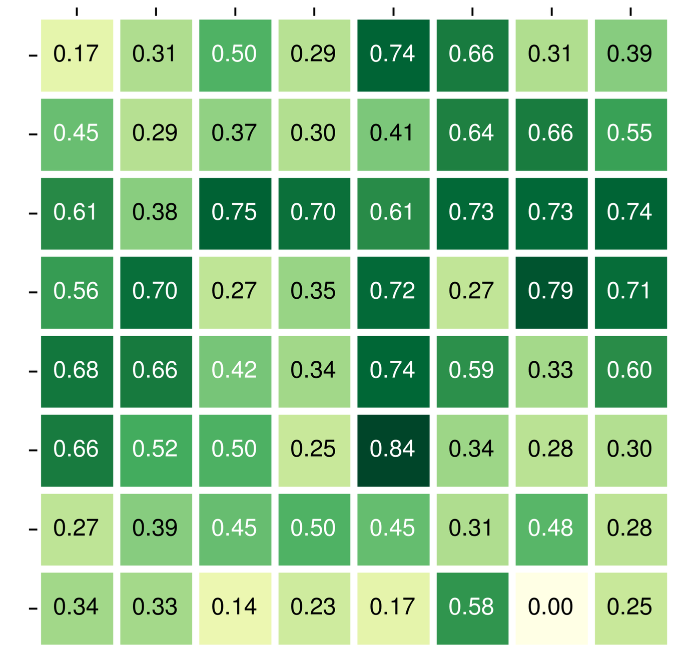
$$[\boldsymbol{\Gamma}_n(\boldsymbol{\theta})]_{ij} = \begin{cases} \sigma_i^2 & \text{if } i = j \\ \sigma_i \sigma_j \rho_{ij}(\ell_1, \ell_2) & \text{if } i \neq j \end{cases},$$

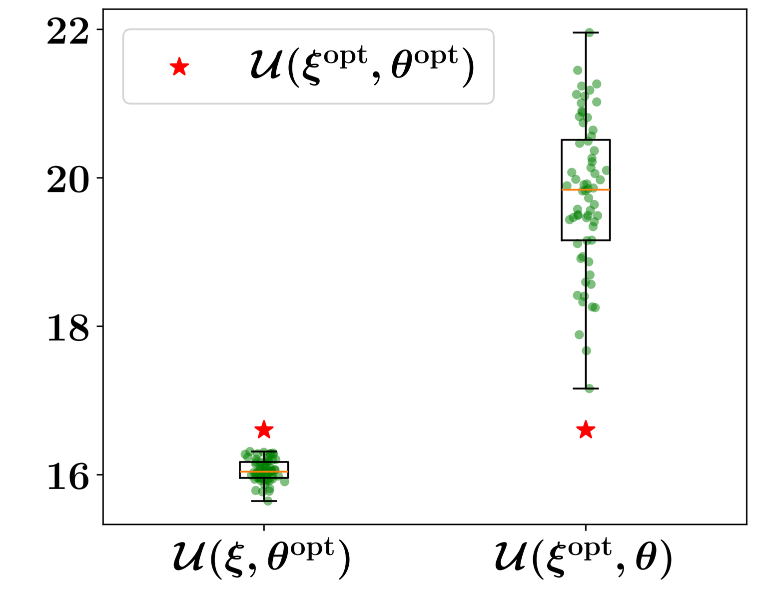
with

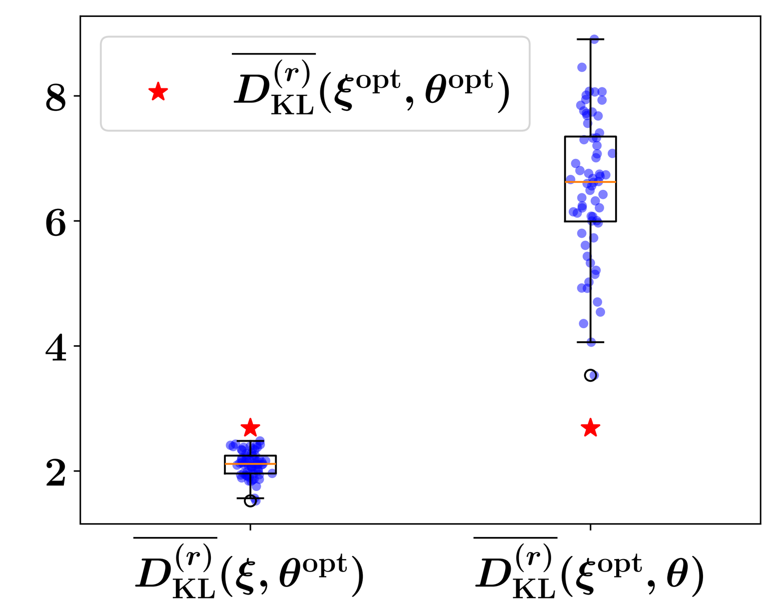
$$\boldsymbol{\theta}_i = \begin{cases} \sigma_i & \text{if } i \leq N_d \\ \ell_1 & \text{if } i = N_d + 1 \\ \ell_2 & \text{if } i = N_d + 2 \end{cases},$$











References

Ahmed Attia.
Probabilistic approach to black-box binary optimization with budget constraints: Application to sensor placement.
arXiv preprint arXiv:2406.05830, 2024.

Abhijit Chowdhary, Ahmed Attia, and Alen Alexanderian.
Robust optimal design of large-scale bayesian nonlinear inverse problems.
arXiv preprint arXiv:2409.09137, 2024.

Abhijit Chowdhary, Shanyin Tong, Georg Stadler, and Alen Alexanderian.
Sensitivity analysis of the information gain in infinite-dimensional bayesian linear inverse problems.
International Journal for Uncertainty Quantification, 14(6), 2024.

Funding

The work of A. Chowdhary and A. Alexanderian was supported in part by US National Science Foundation grants DMS #2111044. The work of A. Alexanderian was also supported in part by US National Science Foundation grants DMS #1745654. The work of Ahmed Attia is supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR), ASCR Applied Mathematics Base Program, and Scientific Discovery through Advanced Computing (SciDAC) Program through the FASTMath Institute under contract number DE-AC02-06CH11357 at Argonne National Laboratory.